DOCUMENT RESUME

ED 122 229 CS 002 585

AUTHOR Summers, Edward G.; Edwards, Peter
TITLE Developing a Canadian Criterion Referenced Core Reading Vocabulary Test for Junior Secondary Grades.
PUB DATE Dec 75
NOTE 83p.

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage
DESCRIPTORS *Criterion Referenced Tests; English (Second Language); Models; Native Speakers; Secondary Education; *Test Construction; *Vocabulary; *Word Lists
IDENTIFIERS *Canada

ABSTRACT

This document describes the development and application of a model for generating a core reading vocabulary screening test based on rank frequency distributions derived from a quarter of a million word corpus of 500 word samples embracing commerce, English, home economics, industrial education, mathematics, science, and social studies. The word selection algorithm is based on probabilistic concepts drawn from work in information science and keyboard indexing and uses frequency data to partition vocabulary lists into sections ranging from high frequency to rare words. Word types selected from these sections constitute general vocabulary, general subject vocabulary, and special subject vocabulary. A high utility 496 word vocabulary list is presented for use in screening both native speakers and second language students who will have difficulty in reading texts assigned in junior secondary subject areas. The word list is cross validated using frequency data drawn from the Kucera and Francis adult word list and the American Heritage elementary list. Suggestions for calibrating and using the core vocabulary test within the emerging framework of criterion referenced testing are also included. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED122229

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Developing a Canadian criterion referenced core reading vocabulary test for junior secondary grades

Edward G. Summers

University of British Columbia

and Peter Edwards

Canberra College of Advanced Education

Faculty of Education

University of British Columbia

Vancouver, B.C.

December 1975

US002585

ERRATA

1. Page 1 footnote

   Change Office of Field Development to Division of Field Development

2. Page 5 - last line prior to quote

   Change Davis (1971) to Davies (1971)

3. Page 16 - line 8

   Change 1975 in Edwards reference to 1974

4. Page 34 - Table 8

   Frequency column, change 6,496 to 6,459; 198 to 197

5. Page 73 - References

   Add: Hunt, B.L., Snyderman, M. & Payne, W. Machine-assisted indexing of scientific research summaries. Journal of the American Society for Information Science, 1975, 26, 230-236.

Developing a Canadian criterion referenced core reading vocabulary test for junior secondary grades*

Edward G. Summers

University of British Columbia

and Peter Edwards

Canberra College of Advanced Education

ABSTRACT

Describes the development and application of a model for generating a core reading vocabulary screening test based on rank frequency distributions derived from a quarter of a million word corpus of 500 word samples embracing commerce, English, home economics, industrial education, mathematics, science and social studies. The word selection algorithm is based on probabilistic concepts drawn from work in information science and keyword indexing and uses frequency data to partition vocabulary lists into sections ranging from high frequency to rare words. Word types selected from these sections constitute General Vocabulary, General Subject Vocabulary and Special Subject Vocabulary. A high utility 496

---

*The data base used in the development of the core reading vocabulary test was derived from the doctoral dissertation of Peter Edwards, "A Computer Generated Corpus and Lexical Analysis of English Language Instructional Materials Prescribed for Use in British Columbia Junior Secondary Grades," University of British Columbia, Vancouver, October, 1974. Computer dollar support from the UBC Faculty of Education and the Computing Centre is also acknowledged. In addition, acknowledgment is made for released time for R & D through the Office of Field Development, Dr. Dave Thomas, Director.

word vocabulary list is presented for use in screening both native speakers and second language students who will have difficulty in reading texts assigned in junior secondary subject areas. The word list is cross validated using frequency data drawn from the Kučera and Francis adult word list and the American Heritage elementary list. Suggestions for calibrating and using the core vocabulary test within the emerging framework of criterion referenced testing are also included.

Introduction

Hundreds of vocabulary studies utilizing word frequency counts have been reported in the literature since the early 1900's. Word frequency lists have been found to be of value in many areas including: language learning, development of standardized tests, determining readability level, selection of textbooks, writing instructional materials, information retrieval indexing, compiling vocabulary lists, collection of data in psychological studies, and in assessing word knowledge in relation to teaching. Thorndike (1921); Thorndike and Lorge (1944); Horn (1926); Buckingham and Dolch (1936); and Dale and Chall (1948) conducted particularly well known studies the results of which have had wide application in education.[1] Recent studies by Kučera and Francis (1967), Carroll et al. (1971), and Harris and Jacobson (1972) are based on more contemporary printed materials. Word studies have been conducted for a great variety of purposes. The one principle underlying them all is that, to a great extent, occurrence frequencies indicate the overall importance of the word.

---

[1]The details underlying these studies have appeared in research studies and reviews countless times and are well known. For an excellent analysis and summary of these and other widely quoted word count studies, see the printed doctoral thesis of Roberts (1965, pp. 9-12), and the introductory sections of the more recent Harris and Jacobson (1972) study.

Word frequencies change over time. The historical period from which frequency data are drawn is important in the use of that data. Thus, interest in and use of the results from newer studies is increasing. Advances in computer technology have also enabled us to manipulate larger quantities of sample data and to examine the results of word count studies in far greater depth. The advantages of computer technology in linguistic research are aptly illustrated by Kučera (1969).

> Since any useful analysis of language usage has to be based on a large body of textual material, even elementary information could be obtained, before the advent of computers, only with enormous labor. Let us imagine that one wished to determine some very basic lexical properties of a textual corpus containing a million running words. If this were to be done by hand (or, more accurately, by the human brain), the task would require an inordinate amount of time; each of the one million words would have to be inspected individually, and each new word recorded after first checking to make sure it had not already been noted. If the analysis were also to preserve information about the frequency of occurrence of individual words, or perhaps references to the pages or lines of the text where their occurrences were to be found, the assignment would become more formidable still. . . . Linguists and lexicographers alike have found in the computer a new and useful tool that has not only made the analysis of language less time-consuming but has also opened new insights into important problems in language usage. (1969, p. xxxviii)

Roberts (1965) makes much the same point.

> Frequency studies of the components of language have always been hampered by the prodigious amount of man-hours required to make the studies extensive enough to be either valid or useful. The great semantic counts and the Lorge 'Magazine Count' of the 1930's were by-products of the depression. For example, The Semantic Count of the 570 Commonest English Words by Irving Lorge employed the efforts of several hundred workers provided by the W.P.A. for a period of six years beginning in 1934. (p. 5)

The current generation of word frequency studies mentioned above would have been impossible without access to the storage capacity of high speed electronic computers and the sophisticated programming techniques available through the present state-of-the-art in computer technology. Within the last two decades computer technology has been revolutionized twenty times over but input (and also output to some extent) is still relatively horse and buggy. A major barrier which still blocks unlimited sampling and processing of natural language text data is input technology. When multi-font automatic optical scanning devices are perfected, sampling and input problems will be of relatively minor concern in language studies by computer.

It must be borne in mind that word lists are never more than representative approximations of the domain they are sampling. The seemingly massive array of word types contained in an unabridged dictionary are only a subset of the total lexicon of English. Thus, the lexicographers dilemma (or salvation)--to be able to deal only with samples and approximations rather than the complete universe of language and meaning. As Davis (1971) points out:

> Dictionaries can still usefully stick to their old task of recording in handy format a small canonical set of attributes for a selection of words, without claiming that the attributes are the totality of lexical information or that the selection of words is a complete listing of lexical items. (1971, p. xii)

Word lists developed for use in education are never more than prescribed subsets of the vast domain of the general universe of English word types. For example. The Kučera and Francis study is based on samples taken from a set of juried materials which were judged to be representative of fifteen genre of adult writing. The resultant word list is only an approximation of the sampled materials and the materials themselves are barely representative of what might be called the infinite corpus of printed American English. The American Heritage Word Frequency Book (Carroll, et. al., 1971) is currently the most carefully developed educationally related word list in print. However, the sample it provides of the printed language of the American elementary educational system is just a subset of the larger lexicon in use in schools. Thus, limitations in the state-of-the-art oblige us to rely on sampling in making word lists. But, if the words in such lists are gathered in controlled fashion, and if they statistically represent a carefully defined subset of the total lexicon, the lists have considerable validity, provide insights into language usage, and can be of great value in teaching and in organizing instructional programs.

Rationale for the Study

The purpose of this paper is to describe the development of a criterion referenced core reading vocabulary list for use in

junior secondary schools. Students arrive at Grade 8 in Canadian schools with considerable diversity in reading ability. Such populations often exhibit reading achievement ranging from primary grades to beginning university level as measured by standardized reading tests. In addition, many students are second language learners with limited reading ability. It is particularly crucial to identify and sort out those students who will have great difficulty in reading the instructional materials of the secondary subject areas so that instruction appropriate to their needs can be organized. A screening test, based on crucial core reading vocabulary, can serve as a useful supplement to informal teacher assessment and the more general standardized reading survey test in identifying such students. An adequate screening test would provide answers to two related questions:

(1) Can the student recognize the most common words which account for a considerable percentage of the natural language appearing in secondary school textbooks?

(2) Can the student recognize the most common special vocabulary which frequently appears in the instructional materials used in the subject areas?

Low scores would indicate minimal development of the functional recognition vocabulary necessary for reading secondary instructional materials. The test should be designed to screen vocabulary needs for two populations--native speakers of English who have reading problems and students learning English as a second

language. In addition, it should be adaptable for use with groups or for individual testing.

In scanning a passage, the skilled reader familiar with the language recognizes the most common words on an almost automatic basis. To him, the special subject words provide the most significant new information in the passage coupled with the word order, structure, and logical and semantic relationships established by the author. In a sense, the special subject words and the common words become informing and non-informing vocabulary in terms of information gain. The passage is redundant, or difficult, depending on the reader's familiarity with the special subject words. This sorting out of connected discourse, involving the redundant and the unfamiliar, both in terms of word knowledge and broader meaning, is the major process in comprehension and is often described linguistically by other authors as the drive toward "reduction of uncertainty" in reading. Reading has also been referred to as a "psycholinguistic guessing game" wherein the reader matches his reservoir of language experience with print and uses cues from both in arriving at meaning. The closer the match between the semantic and syntactic structure of the passage and the linguistic knowledge of the reader, the higher the comprehension. The reader's apprehension of the subject matter of a passage is obviously determined by his ability to recognize the common and special subject words making up the passage.

To the neophyte reader very little on the printed page is redundant. Every graphic character represents a difficult recognition challenge. The older reader who does not automatically read and understand the most common words, which constitute a substantial percentage of the printed page, will be hopelessly lost when faced with the diverse reading demands of the secondary curriculum. The problem is particularly acute for second language learners. In learning English, the "stream of speech" takes precedence and the instructional tasks move from the sound system, to structure, to vocabulary and then to reading and writing. Good development in speaking and understanding is necessary before moving to reading. However, many secondary students are faced with the problem of coping with the instructional materials of the program with only limited speaking and understanding knowledge of English and little reading ability. A core vocabulary reading test, that is representative of the functional lexicon of secondary instructional materials, can be useful in determining if such students possess enough reading vocabulary to profit from instruction using secondary textbooks.

Alford (1971) suggests that the person who can recognize the 1,000 most common words in the language will, on the average, meet only one or two new words per sentence. There is considerable probability that the learner can deduce their meaning from context. When 3,000 high frequency words are known, only about

one word in a paragraph presents a comprehension problem. Such words constitute a large part of the structural operation of English and the vocabulary which serves as the basic threshold for entry into subject materials. Obviously, it is reasonable to first teach the learner the most frequently used part of the language and provide a solid base for later extension into more specialized materials. All language development programs would agree on this whether the word lists used are based on frequency counts or some logical principle of selection.

In word frequency studies, language can be defined as a finite or infinite set of sentences, each finite in length and concatenated from a set of finite elements. The finite set of elements consists of the alphabet and the vocabulary of the language. Of course the elements must be joined in something other than haphazard fashion. Criteria for forming the vocabulary into meaningful sentences are contained in the grammar of the language. The grammar is, in essence, the theory of the language. Word frequency lists are aids in focusing on the vocabulary elements. The typical algorithm used in developing word lists consists of the following steps.

(1) Define the subset of language which is of interest.

(2) Select texts and develop a corpus of materials to represent the subset.

(3) Using sampling techniques, generate a body of passages which

represent the corpus of materials.

(4) Produce lists of all the words contained in the passages and indicate the frequency of occurrence for each word type.

(5) Organize the word types into an alphabetical list and a list ranked in order of decreasing values of F.[1]

(6) Organize the lists on other dimensions such as subject or grade level.

Core vocabulary can then be selected from the ranked frequency lists. The fundamental question is one of deciding which words from the ranked list serve as the best indicators of core vocabulary. Frequency lists often run to thousands of word types. How does one decide which words convey the most information and best represent the actual subject content of the passages? How many words should be selected? Is there a systematic method for partitioning the word frequency list and selecting the best class of words? Is the method based on a definable theory of word significance? Obviously, some words are good core words because of their high frequency, others are good because of their lexical value, while others are poor because of their rarity. No completely adequate theory of word significance has yet been put forth for use in vocabulary studies. We would opt for frequency

---

[1]Other ranking strategies are often employed. A common one is to rank word types according to decreasing values of the ratio of the relative occurrence frequency for a particular word to the relative occurrence frequency for all words in a corpus. Keyword indexing schemes use various relative frequency rankings in selecting terms for information retrieval. (See Salton, 1968, pp. 359-362 and Sparck Jones, 1971).

of occurrence, coupled with subjective judgment, as the most useful indicator of word significance.

Frequency is a particularly good measure in selecting core words for screening purposes. Frequently occurring words provide the best clues to information content and pinpoint the subject category to which the materials belong. Frequency is also a good indicator of the likelihood that a language user will be familiar with a word and its meaning. Thus, we would suggest that core vocabulary be drawn from the high frequency sections of word lists. However, we would also emphasize that the extraction of words from frequency lists cannot be done solely on a statistical basis but also involves considerable subjective judgment based on semantic considerations.

Although word frequency studies have been reported based on materials used by children and adults, few studies have utilized samples from secondary materials, particularly those prescribed for Canadian schools. This study provides a core list derived from textbooks representative of those most students come in contact with in junior secondary grades in British Columbia schools. The list is also validated using other recent word frequency studies based on elementary and adult materials. The core vocabulary can be useful in assessing and teaching older students and adult learners experiencing difficulty in reading. It also has some potential use in upper elementary grades and perhaps in special

education and learning disability programs. In addition, the lists could be particularly useful for screening and teaching secondary students and adults learning English as a second language and attempting to cope with the reading demands of subject content. There is increasing need for placement tests for use with this population. Recent estimates suggest that 20 to 25 percent of British Columbia secondary students are coping with English as a second language.

In the following sections, we describe the selection and processing of the basic corpus of junior secondary instructional materials; analyze the word frequency results; discuss the rationale used in partitioning the frequency list and selecting core vocabulary; and present the final vocabulary list. In addition, suggestions for using the list are given within the framework of criterion referenced testing.

The Basic Corpus of Materials

The basic corpus consisted of 469 text samples of approximately 500 words each (or as near 500 as possible and retain complete sentences). The text samples were systematically stratified to represent the 37 A issue texts prescribed for use in Grades 8, 9, and 10 for Commerce, English, Home Economics, Industrial Arts, Mathematics, Science, and Social Studies (British Columbia Department of Education, 1972). Table 1 presents the number of texts and samples for each grade level and subject area.

INSERT TABLE 1 ABOUT HERE

Eleven texts were not included because of their especial characteristics and the difficulty of obtaining adequate samples of natural language prose. These included 2 Commerce texts, 4 English texts, 2 Mathematics texts, and 3 Social Studies texts. The above contain large amounts of shorthand exercises, poetry and blank verse, algebraic and geometric problems, and atlas content. Materials assigned to one grade and repeated in another were sampled only once (total of 9 texts). The above procedures resulted in a sample of approximately 500 words being selected from roughly every 20 pages throughout the 37 junior secondary textbooks. The basic corpus represents a wide variety of language and consists of approximately 40 to 50 percent of the total natural language prose presented in the 37 textbooks.

Word Frequency Results

Standard input rules were followed, with words defined in terms of graphic form, and the 469 text samples keypunched and transferred to magnetic tape for computer processing. Use of several unique machine based editing procedures, along with considerable hand editing, resulted in an almost nil error rate in replicating the samples for computer analysis. Programs were developed to produce hard copy of the samples and to generate

TABLE 1

Number of texts and samples by grade level and subject

Subject	Grade 8		Grade 9		Grade 10		Sub Total	
	Text	Sample	Text	Sample	Text	Sample	Text	Sample
Commerce	X	X	2	25	2	16	4	41
English	2	17	4	47	2	16	8	80
Home Economics	1	22	5	76	X	X	6	98
Industrial Ed.	1	9	3	54	X	X	4	63
Mathematics	1	14	1	7	1	14	3	35
Science	2	20	2	24	2	31	6	75
Social Studies	2	22	2	13	2	42	6	77
Grade Totals	9	104	19	246	9	119	37	469

ranked and alphabetical frequency lists including:

(1) total corpus of samples (2 lists)

(2) samples organized by grade level (6 lists)

(3) samples organized by subjects across grades (14 lists)

(4) samples organized by subjects within grades (36 lists)

(5) samples organized by textbooks (74) lists) (Miller, 1974)

In addition, X, FX SUM FX, FX* X AND CUM% FX *X statistics were generated for each of the 66 ranked frequency lists. (Edwards, 1975, p. 84)

A suite of programs was also developed to process the samples allowing for descriptive and comparative analyses on various facets of the corpus including measures of central tendency, variability, word and sentence repeat rate frequency and word and sentence length distribution. Analysis of results using these data will be included in a separate report now in preparation. Table 2 presents the type, token, character and word length characteristics for the total corpus, each of the three grade levels, and the seven subjects across the grade level samples. The total corpus of 469 text samples produced 235,107 running words (tokens) and 16,405 distinct words (types).

INSERT TABLE 2 ABOUT HERE

TABLE 2

Number of types, tokens, characters, and average word length in characters for the corpus, grade levels and the seven subjects across grades

	TYPES	TOKENS	CHARACTERS	AVERAGE
Corpus	16,405	235,107	1,062,411	4.52
Grade 8	7,027	52,867	234,527	4.44
Grade 9	11,401	122,953	554,488	4.51
Grade 10	7,736	59,343	273,654	4.61
Commerce	3,020	20,153	90,171	4.47
English	7,079	40,300	178,192	4.42
Home Economics	5,529	49,257	221,576	4.50
Industrial Ed.	4,060	31,300	141,176	4.51
Mathematics	1,952	17,808	73,852	4.15
Science	4,833	37,787	173,023	4.58
Social Studies	6,211	38,608	184,727	4.78

The word types identified for each of the seven subject areas range from a low of 1,952 for Mathematics to a high of 7,079 for English. These are rather interesting results. The texts sampled do contain a good percentage of special subject words. Other studies suggest that specialist subject vocabularies generally run to about 10,000 words. This vocabulary would encompass a good percentage of the most used special subject lexicon. Subject vocabularies do differ markedly from each other both in terms of content and frequency of occurrence of word types. However, one feature remains fairly constant. A small number of frequently occurring word types will account for a large percentage of the total running prose of the materials sampled. Another way of putting this is that we can expect a shortage of words with very high occurrence frequencies and an excess of words with very low occurrence frequencies. This is well illustrated by the results in Table 3 where the percentages of total text accounted for by the 100 most frequently occurring word types for the samples organized by corpus, grade levels and subjects are compared. For the total corpus, the 100 words with the highest frequency of occurrence account for almost 50 percent of the running text in the 469 samples used. The results are extremely uniform across the samples organized by grade and subject areas. As a sidelight, we should point out that, interestingly enough, this pattern also obtains when abstracts of documents are analyzed. Hunt, Snyderman

INSERT TABLE 3 ABOUT HERE

and Payne (1975) recently reported data on this. With a sample of 3,445 documents, it was found that 27 words accounted for 33.3 percent of word usage in abstracts. These included: the, of, and, to, in, a, be, will, for, on, is, with, by, as, this, are, from, that, or, at, been, an, was, were, have, has, it. The results are very typical of those obtained in frequency studies based on document samples of varying size.

Table 4 compares the overall word frequency results with those from other studies using the parameters of source of sample, size of sample, number of samples, number of tokens and distinct word types.

INSERT TABLE 4 ABOUT HERE

The 16,405 word types is about the volume of vocabulary output one would expect, given the number of samples and sample sizes, in comparison with the much larger bodies of text used in the Kucera and Francis, Carroll et al. and Harris and Jacobson studies.

Generating the Core Vocabulary List

In defining a core vocabulary, the basic problem consists of selecting the most significant words from frequency ordered

TABLE 3

Percent of total text accounted for by the 100 most frequently occurring words

Corpus	48.97 %
Grade 8	50. 59
Grade 9	49.17
Grade 10	48.28
Commerce	52.87
English	50.93
Home Economics	52.25
Industrial Educ.	50.69
Mathematics	59.48
Science	59.12
Social Studies	48.05

TABLE 4

Comparison of four word lists on source of samples, size of samples, number of samples, number of tokens, and number of distinct word types

Word List and Source	No. of Samples	Size of Samples	No. of Tokens	No. of Word Types
Kučera and Francis (1967)				
samples from adult materials from 15 genres of writing	500	2,000 words	1,014,232	50,406
Carroll et al. (1971)				
samples from 22 categories of curriculum materials used in grades 3 through 9 representing 6,162 school textbooks	10,000	500 words	5,088,721	86,741
Harris and Jacobson (1972)				
127 books representing 14 series of ele. texts (grades 1-6) including 6 basal reading series and two series each in math, Eng, social studies and science	publishers word lists or every word in the series		4,500,000	80,000
Summers and Edwards (1975)				
37 prescribed texts in commerce, English, home economics, industrial arts, soc. studies, science and mathematics in grades 8, 9 and 10	469	500 words	235,107	16,405

vocabularies based on an adequate number of sample passages which have been carefully stratified to represent a wide variety of secondary textbooks. The critical question is, of course, What constitutes core vocabulary? The same question is also frequently posed in the field of information science where the problem is one of selecting keywords to serve as surrogate representations of documents in indexing for information retrieval purposes (American Society for Information Science, 1975). Work in this area has stemmed from Luhn's (1958) earlier studies utilizing occurrence frequency statistics in keyword identification. The concepts related to keyword selection are also extremely relevant in attacking our problem of selecting core vocabulary from ranked frequency lists.

The Summers and Edwards study produced a ranked word frequency list based on the total corpus of samples from the 37 subject textbooks. Table 5 presents the number of repetitions for each occurrence frequency for the total corpus ranked in descending order. At the top of the distribution read, one word type had 17,519 occurrences, one word type had 8,177 occurrences and so on. At the bottom read, 7,098 word types occurred only once, 2,418 word types occurred only twice and so on. The first

---

INSERT TABLE 5 ABOUT HERE

TABLE 5

Repetitions for each occurrence frequency for the Summers and Edwards corpus

Tokens	Types	Tokens	Types	Tokens	Types	Tokens	Types
17519	1	663	1	363	1	200	3
8177	1	639	1	358	1	197	1
6459	1	581	1	353	1	196	1
5022	1	569	1	351	1	193	1
3913	1	558	1	349	1	191	1
2266	1	552	1	342	1	190	1
2218	1	542	1	340	1	189	3
2165	1	537	1	312	1	184	1
2141	1	529	1	309	1	183	1
2090	1	518	1	306	1	182	2
1871	1	510	1	303	1	179	1
1789	1	498	1	295	1	178	3
1650	1	491	1	286	1	177	1
1485	1	488	1	279	1	174	1
1459	1	481	1	268	1	171	1
1291	1	476	1	255	1	167	1
1246	1	465	1	254	1	166	1
1110	1	463	1	253	1	164	1
1087	1	439	1	245	1	163	1
1064	1	428	1	239	1	162	2
1057	1	425	1	236	1	156	1
1054	1	424	1	234	2	155	1
992	1	423	1	233	1	153	3
933	1	421	1	232	2	151	2
889	1	417	1	228	1	150	1
857	1	411	1	223	1	149	2
856	1	406	1	220	1	148	3
833	1	403	1	217	1	147	1
826	1	400	2	216	2	146	1
816	1	393	1	215	1	143	2
779	1	386	1	211	1	141	1
772	1	369	1	208	1	140	2
700	1	368	1	205	1	138	4
681	1	366	1	202	1	136	2

TABLE 5 (continued)

Tokens	Types	Tokens	Types	Tokens	Types	Tokens	Types
135	5	95	2	60	5	26	44
134	1	94	5	59	6	25	37
133	1	93	3	58	9	24	44
130	3	92	4	57	8	23	36
129	2	90	4	56	2	22	63
128	2	89	2	55	9	21	51
127	1	88	5	54	18	20	68
125	1	87	3	53	10	19	58
124	2	86	2	52	8	18	98
123	1	85	2	51	13	17	98
122	2	84	4	50	14	16	100
121	3	83	2	49	8	15	113
120	2	82	6	48	12	14	138
119	2	81	5	47	10	13	123
118	4	80	4	46	18	12	156
117	3	79	9	45	13	11	189
116	1	78	5	44	17	10	208
115	1	77	6	43	24	9	251
114	3	76	5	42	21	8	322
112	1	75	7	41	18	7	387
111	2	74	4	40	26	6	453
110	2	73	10	39	21	5	632
109	8	72	8	38	25	4	854
107	3	71	5	37	26	3	1240
105	2	70	4	36	31	2	2418
104	4	69	10	35	40	1	7098
103	4	68	10	34	39		
102	2	67	11	33	21		
101	3	66	10	32	30		
100	3	65	4	31	27		
99	3	64	7	30	38		
98	2	63	13	29	33		
97	4	62	12	28	34		
96	3	61	3	27	39		

six word types in the distribution account for approximately 20 percent of the total tokens in the corpus. Extending this to the first 100 most frequent word types, we find that these account for approximately 49 percent of the total tokens. These few syntactical words appear hundreds of times throughout the samples taken from the 37 textbooks. At the lower end of the distribution, the 12,695 word types occurring six times or less account for only 10 percent of the total tokens in the corpus. This end of the distribution contains several thousand words that appear very infrequently.[2] The 100 most frequent words, and the 12,695 words which appear six times or less, thus account for 78 percent of the total word types (16,405 words) contained in the 469 samples from the 37 subject textbooks. Between these two extremes fall the remaining 22 percent of word types in the total corpus (3,709 words in all). It is our contention that the most useful special subject vocabulary lies within this relatively small group of words.

The distribution of 267 occurrence frequencies given in

---

[2]It is difficult to establish anything but a subjective rationale for this cutoff point. For the corpus as a whole, occurrence frequencies of six times or less seemed a good point for establishing word utility whereas in the vocabularies for the seven individual subject corpera, a cutoff of three or four seemed more appropriate. These cutoffs also reflect those established in other studies. An alternative technique would be to specify that a given term has to constitute X percentage of the total text of the sample to be considered significant, a more difficult estimate than the above procedure.

Table 5 provide graphic illustration of the plight of the secondary student with poorly developed reading skills. Of the total vocabulary of 16,405 words, hundreds occur only six times or less throughout the samples from the 37 textbooks. This is reading load enough but the situation is exacerbated even more when we consider that most secondary classes also use a wide variety of reference and supplementary materials, in addition to the standard textbooks, thus increasing the vocabulary load even more. These supplementary materials tend to be even less uniform in their vocabulary control. On the other hand, a small number of subject words, and a set of high utility common words, will appear with considerable regularity. Obviously, this core functional vocabulary is the key element in coping with secondary textbook reading demands. A screening test, designed to assess student knowledge of this vocabulary, could be an extremely useful diagnostic and instructional tool.

Before moving to discussion of techniques for selecting core vocabulary, we will digress briefly and examine another intriguing question. Subjects will differ in the sheer quantity of words used and all exhibit a varied range of associated vocabulary demands. However, do subject areas differ in the repetitiveness of their vocabularies? Yule's (1944) K characteristic was one of the earliest vocabulary statistics and provides a measure of the concentration of vocabulary in samples of

natural language text. Most importantly, the K characteristic is independent of sample size when they are randomly selected and sufficiently spread across the population being studied (Yule, p. 44) which allows for comparison of various size samples of text. K is, in essence, a measure of the probability that a randomly chosen pair of words will be identical. The larger the value of K the greater this probability. By way of illustration, Yule estimated K for St. John's Gospel to be 177.9 while it was 56.5, 66.9, 80.6, and 88.0 for four selected works of John Bunyan. (Note: K can be computed for all words, nouns only, or any aspect of literary style thought to be of significance).

Ranked K values for the 37 textbooks for the Summers and Edwards corpus are shown in Table 6. The corpus K value of

---

INSERT TABLE 6 ABOUT HERE

108.104 can be used as a basis for comparison. Five of the Home Economics texts (77 samples) tend to have the lowest repeat rate for vocabulary while the six Social Studies texts (77 samples) have the highest repeat rate. It is interesting to note the split in the English texts. The prose, short story and reading instruction texts (6, 8, 9, 10, 11) have low repeat rate frequencies while those texts geared to language instruction *per se* (5, 7, 12) tend to have a greater repetition rate. It is difficult to

TABLE 6

Ranked K values for 37 subject textbooks

Subject		Ranked K value	No. samples
Home Economics	17[a]	81.857	10
Home Economics	16	87.723	14
Home Economics	18	92.203	9
Commerce	3	95.532	7
Home Economics	15	97.723	22
Home Economics	13	98.166	22
English	9	99.231	10
English	10	99.253	10
English	6	101.873	10
English	11	102.632	12
English	8	103.655	7
English	7	105.651	20
Home Economics	14	111.747	21
Commerce	4	111.939	9
Industrial Education	20	113.084	13
Industrial Education	21	114.428	16
Commerce	1	114.560	11
Industrial Education	19	116.973	9
Science	27	117.664	11
Science	31	117.712	14
English	5	118.117	7
Mathematics	24	118.672	7
Mathematics	23	123.568	14
English	12	125.065	4
Commerce	2	125.988	14
Social Studies	34	126.723	8
Social Studies	33	127.347	7
Social Studies	37	128.258	12
Science	30	128.905	17
Mathematics	25	131.571	14
Social Studies	35	134.655	5
Social Studies	32	137.026	15
Social Studies	36	137.962	30
Science	29	142.048	11
Science	28	150.283	13
Science	26	167.198	9
Industrial Education	22	169.462	25

[a]Numbers refer to specific textbooks listed in the reference section

state, however, that the vocabulary of Home Economics is more or less difficult than that of Social Studies. Qualitative judgments would have to enter here in terms of the semantics involved. Theoretically, a subject could very well have a high repeat rate but the words themselves could be extremely rare and difficult. Thus, word potency also enters in. However, the K characteristic does provide broad differentiation in terms of sheer repetition of words aside from the semantic aspects and K forms a basis for instructional considerations in text comparison.[3]

The Kučera and Francis study based on 15 adult genres of writing provides the closest comparison for these repeat rate results (Table 7). The overall differences in magnitude may be

---

INSERT TABLE 7 ABOUT HERE

---

accounted for by the fact that we are dealing with two essentially different subsets of language. For example. The median sentence lengths for the Kučera and Francis study range from a low of 9.98 words for the Fiction: Mystery Detective genre to a high of 21.97

---

[3]The K factor results were intuitively demonstrated during the actual selection of core vocabulary words from the ranked distribution lists for the seven subject areas. Special subject words of high frequency were much less obvious in English and Home Economics, with occurrences tending toward a flat distribution, while in the other subject areas they "peaked" and stood out more dramatically.

TABLE 7

Selected K factors Kučera and Francis and
Summers and Edwards corpera

	Highest K	Corpus K	Lowest K
Kučera and Francis	402.328 (Fiction: Mystery Detective)	266.240	217.739 (Misc: Govt. Publications)
Summers and Edwards	130.372 (Mathematics)	108.104	92.572 (Social Studies)

words for the Miscellaneous: Government Publications genre with a corpus median sentence length of 16.41. The standard deviations are 9.45, 16.25, and 12.37 respectively. The median sentence lengths for the Summers and Edwards corpus range from a low of 13.53 words for the Mathematics samples to a high of 18.20 words for the Social Studies samples with a corpus median of 15.74. The standard deviations are 8.15, 9.58 and 10.05 respectively. Thus, the Kučera and Francis corpus has a broader base of samples, more genre variety and most likely significantly longer sentences and greater variability in sentence length. In general, the K factor differentiates among samples of writing in both subsets of language and the Summers and Edwards results tend to parallel those found for the Kučera and Francis corpus.

Many variables have an influence on comprehensibility of print. Bormuth (1969) identified a total of 169 such variables, 94 of which correlated significantly with passage difficulty. Redundancy based on word repetition is undoubtedly related to comprehensibility and estimates of repeat rate frequency can provide additional insight into the vocabulary demands of print materials. Students with minimal reading development will have greater difficulty in reading instructional materials with uncontrolled vocabulary repetition. It is particularly important that repeat rate be considered in writing the textbook materials that provide initial entry into the fundamental ideas and concepts of

the subject area. Careful vocabulary control and repetition has long been a hallmark of elementary instructional materials but tends to receive less emphasis at the secondary level. (Fry (1972) provides a good discussion on elementary vocabulary control and Pennock (1974) presents a quick word list for B.C. elementary grades).

We'll now return to the problem of developing core vocabulary lists. Selection of the most useful high frequency words for a core vocabulary is relatively straightforward. However, selecting the most useful special subject words is another matter. Again, keyword indexing studies from information science provide useful insights. Harter's (1975a) recent paper cites previous efforts in automatic keyword indexing and reviews studies by Stone and Rubinoff (1968); Maron (1961); Curtice and Jones (1967); Bookstein and Swanson (1974); Bonwit and Aste-Tonsmann (1970); and Dennis (1965). Harter's research focuses on the statistical procedures used in identifying the technical vocabulary characteristics of a body of literature. The task was to devise a method of automatically selecting terms from abstracts to act as index tags for documents. An experimental collection consisting of 650 abstracts (mean length 223 words) on the works of Sigmund Freud served as the data base. The results suggest that special subject words can be selected from non-special words based on the assumption that deviation from randomness is indicative of language

structure. Harter presents this rationale:

> . . . consider the two words cathexis and see. It is an empirical fact that, with respect to the experimental document collection occurrences of cathexis tend to be found together in the same documents. To observe this empirical fact is to discover a certain order and regularity in the language of psychoanalysis. A document containing one occurrence (token) of cathexis is likely to contain more than one token of cathexis; occurrences of cathexis are not statistically independent. Conversely, there is no tendency for occurrences of the word see to occur in the same documents; the distribution of see in the document collection displays a lack of order and structure. Tokens of see apparently occur independently of one another, and in this sense appear to be distributed at random throughout the collection of documents. (1975, p. 198)

Thus, randomly distributed words whose frequency distributions approximate a Poisson distribution (positively skewed, leptokurtic) are likely to be the non-specialty words. Harter's concept can be illustrated by using a selected list of 38 words drawn from the ranked frequency distribution for the Summers and Edwards corpus (Table 8).[4]

INSERT TABLE 8 ABOUT HERE

The distributions for the very common words the, and, as and all are similar to Harter's example of see and are obvicusly non-

---

[4]Although the occurrence frequencies are grouped by subject and textbooks comprising subjects, the pattern of occurrence is very similar to that which obtains when one views a breakdown of the 469 samples grouped singly.

Table 8 Occurrence frequencies for 38 selected word types

	Commerce texts				English texts								Home Economics texts						Ind Educ texts				Math texts			Science texts						Social Studies texts						frequency
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	
the	407	578	182	296	263	355	677	253	319	326	460	146	600	644	602	325	242	239	336	447	585	1419	468	275	520	452	431	643	510	716	514	681	273	363	217	1297	458	17,519
and	114	174	72	111	103	146	285	82	213	152	211	65	365	354	358	236	153	187	152	189	217	297	130	58	99	72	108	147	99	188	132	268	118	147	77	393	187	6,496
as	39	48	16	32	36	33	81	20	38	40	35	26	110	117	63	55	31	30	39	38	48	89	47	37	64	29	38	47	31	69	60	44	29	26	21	148	35	1,789
all	19	31	21	15	5	21	36	13	20	9	30	2	30	44	25	22	21	17	12	12	21	16	23	2	24	5	12	16	23	19	11	19	17	13	6	31	18	681
water		1	5		1	5	4	5	1	5	8	1	3	18	13	2	3	12	5	12	3	14		5		6	9	81	13	16	21	21	2	4	2	36	3	340
different	2	8	1	2	2	1	10			1	1		11	23	6	7		2	9	1	11	9	8	8	8	3	17	10	8	8	5	3	3		10			198
surface	1				6	2	1						5	2	1			1	5	4	7	15		4		10	9	11	17	15	2	1	1	3		12		135
world	5		3	1	2	1	7	4	4		2	3																									32
family	1	1	4				2		1				35	29	4	18	4	6											1			3	2	3			4	118
language	2		1		2	1	54	1	1										3	1			3										2			4		75
author						1		1	3																												1	6
beverages														1				9																				10
calories														2				9																				11
canned				1									11			1		3																				16
chamfer																				9																		9
cheque		14																																				14
chloride																										1				2	7	1						11
clerk		11		10	1																																	22
congruent																								7														7
credit	1	8	3	13																																		25
customer		8		9																																		17
entertaining													3	6				1																				10
factory		6		1																																		7
intestine																										1		9										10
joints																			2	9																		11
menu													5	3				3																				11
merchandise		7		2																																		9
nuclei																												4	5	1								10
numeral																							3	1	6													10
plywood																				7	3																	10
prairie																																				6	2	8
protons																												9		1								10
rainfall																															7					1		8
resources																															1		1			4	1	7
salad													3	10				1																				14
singular							8																															8
theatre											1	6	1																									8
throttle																						10																10

specialty words.[5] Tokens for these words tend to occur independently of each other and their distributions would be essentially Poisson in nature with samples in the collection equally likely to contain occurrences. We would label these General Vocabulary and divide them into two categories, very high frequency words and relatively high frequency words. The word types water, different, surface, world, family and language also illustrate Harter's rationale. However, we would not strictly label these specialty words. They are likely to cluster in the samples from a particular subject but also tend to be found in varying degree in samples from other subjects. Thus, in terms of core vocabulary, we would place them in a category between specialty and non-specialty words and label it General Subject Vocabulary.

The balance of the list consists of specialty words and we would place them in a category labelled Subject Vocabulary. These words tend to be found together in samples from texts in the same subject and, according to Harter, do not occur as statistically independent events. The Poisson model would not describe their distributions. Harter statistically confirmed the above intuitive impressions for a small sample of words based on abstracts of documents from the psychoanalytic literature. Although

---

[5] Harter did not analyze the obvious non-specialty words which comprise the category often called common words. However, these are important words and for our purposes would be included in a core vocabulary list.

we did not run similar tests, there is little reason to believe the notion would not also be confirmed using our data based on natural language samples from the texts in the Summers and Edwards corpus. Harter also describes development of a mathematical two Poisson model which proved to be relatively effective in separating specialty from non-specialty words. His 1975b paper goes one step further and outlines an effective probabilistic model for keyword indexing. The algorithmic approach defines a measure of indexibility which reflects the relative significance of words in documents.

A general model for selecting core vocabulary thus would give low significance to rare words. High significance would be accorded to frequently occurring special subject words, what we have labelled general subject vocabulary, and common words with high or relatively high occurrence frequencies. In keyword indexing, specialty vocabulary is normally emphasized with low importance attached to infrequent words and the highly common words. It is helpful to plot the word frequency data presented in Table 5 as an aid in visualizing selection of core vocabulary. Plotted word frequency data typically assumes the general shape presented in Figure 1 (Zipf, 1945, 1949; Mandelbrot, 1953, 1966;

INSERT FIGURE 1 ABOUT HERE

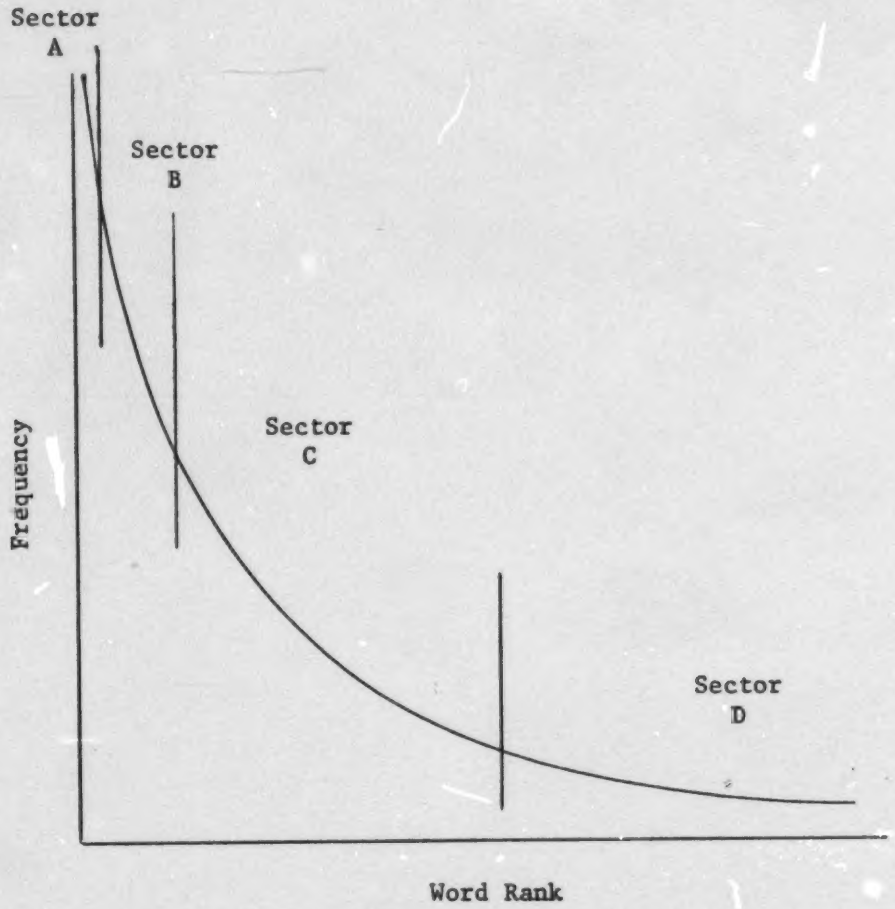

Figure 1. Relationship of word significance to frequency

Herdan, 1956; Meadow, 1973).[6] The abcissa represents the individual word types ranked in order of frequency of occurrence.

The word frequency curve can be partitioned as an aid in selecting core words. The partitioning concept was advanced in early papers by Luhn (1958) and Maron (1961) in their work related to the notion of probabilistic indexing and the development of techniques for machine based abstracting and indexing. Meadow's (1973) chapter on "Statistical Selection of Index Terms" carries the concept further and incorporates it in a discussion of the development of a theory of word significance related to automatic abstracting, indexing and retrieval. Salton, Yang and Yu's (1975) discussion of the discrimination value of words is also relevant here as is the later work on formal decision theory in indexing by Bookstein and Swanson (1975).

---

[6]The "Zipf" curve and its later modifications illustrates the relationship between rank and frequency and is often given as a linear equation represented by a straight line in a two dimensional graph (log rank and log frequency). This is particularly useful as rank-frequency curves from various sources can be compared by inspecting the slope of the lines. A good deal has been written and argued about the Zipf distribution and the wide array of phenomena which it apparently fits. His 1949 text postulates that because of the "law of least effort" words are shortening over time. His 1935, The Psycho-Biology of Language: An Introduction to Dynamic Philology, was offered in a reprint in 1965. A good future project would be to compare the vocabularies of the seven subject corpera by plotting the slopes of the rank frequency curves and comparing their variability. The recently reported work by Wyllys (1975) provides an excellent model. He reports that significant, detectable variations in rank-frequency curves do exist and appear to have the potential for subject field and even sub-field discrimination in terms of vocabulary.

Sector A. This sector includes the 12 to 15 words which invariably occur with extremely high frequency in word lists. Such every-day function words as the, of, and, a, to, in, is, that, it, are, for, you, be, as, and like make up 25 percent of the total words encountered in most natural language prose. Alford (1971) provides a typical breakdown of high frequency words in tabular form. The appropriate figures for the Summers and Edwards corpus have been added in Table 9 for purposes of comparison. Again, a small number of word types invariably

---

INSERT TABLE 9 ABOUT HERE

---

account for a large percentage of running words and a relatively small vocabulary represents a good deal of the total volume of tokens in print.

Sector B. Words in Sector B also occur with very high frequency but with somewhat less regularity than those in A. The combined non-specialty words from A and B constitute anywhere from 30 to 50 percent of the total words encountered in natural language prose. These are the words that would be considered "noise" in information retrieval systems (a la Harter) and would be eliminated by "stop" lists such as those described by Salton (1968). These words would "retrieve" every document if used in information retrieval indexing. The words are often referred to

TABLE 9

Text coverage of high frequency words[a]

No. of Different Words	Approximate % of Text Coverage	
	Alford	Summers and Edwards
4	10	16
50	34	42
100	45	49
350	70	62
1,000	86	74
2,750	96	86

[a]Alford figures based on Russian physics data

as common words in English or as sight or service words in reading and consist mostly of the function words (Fries, 1952; Lefevre, 1964; Goodman, et al., 1966). Fries identifies the four traditional grammatical classes of words (nouns, verbs, adjectives, adverbs). To these he adds a fifth class (conjunctions, prepositions, pronouns, articles) of 154 words, in 15 categories, which serve as structural syntactic markers in a sentence to express relations of grammatical structure.

Present day English has developed wide use of function words and they have a great variety of meanings depending on how they are used. For example. Nine prepositions out of 30 account for 90 percent of the occurrence of prepositions in English. The dictionary lists an exceedingly large number of separate senses for each ranging from a high of 63 for of to a low of 15 for from and a mean of 36 for all nine (Fries, 1940). Function words are important as non-specialty core vocabulary because of their dense distribution in natural language text. Fries found that almost a third of a sample of 1,000 words were function words and many of the most commonly occurring words in the Kučera and Francis (1967) and the Carroll et al. (1971) corpera are function words. (200 function words comprise 50% of the AH corpus, p. xlii).

Function words are often referred to as non information bearing words although we should not dismiss them too lightly for

their semantic contribution. They play an important role in pedagogy, dictionary development, authorship attribution studies and cryptanalysis. In an often quoted paper, Mosteller and Wallace (1963) were able to resolve the disputed authorship of the Federalist Papers by analyzing word distribution and showed that function words exhibit a surprising power to discriminate among authors. Subjects will vary even in the distribution and use of the most common words. The results of the Kučera and Francis (1967) study suggest marked differences in the use and distribution of the 100 most frequently used words in the corpus across 15 different genre of adult materials sampled. Chi-square results from the study reported here also suggest wide author variation in style. The samples from the seven subject areas differed significantly in the distribution of the 100 most common words from the total corpus even when adjustments were made for variation in the amount of material sampled. (A far ranging set of papers on the use of statistical analysis in stylistic studies by Doležel & Bailey (1969) should be noted. Bailey's chapter on "Statistics and Style: A Historical Survey" is particularly useful). Style can also be viewed as a probabilistic concept although those in the field of stylistics and computational linguistics would readily agree that we are still far from development of an adequate theoretical framework for the statistical analysis of stylistics.

Young (1973) emphasized the important role function words play in providing structural information and elminating ambiguity. He illustrates the point by quoting from Lewis Carroll's poem "Jabberwocky":

> Twas brillig, and the slithy toves
>
> Did gyre and gimble in the wabe;
>
> All mimsy were the borogoves,
>
> And the mome raths outgrabe . . .
>
> Alice says that 'Somehow, it seems to fill my head with ideas--only I don't exactly know what they are!' The ideas which Alice derives from the poem are generated by the structural patterns and the underscored function words. (p. 69)

Function words are not necessarily "noise in the system," particularly to the neophyte reader or new language learner. Along with form and word order they constitute the grammar of English and add considerably to meaning.

Sector C. This sector contains additional general vocabulary and words which, although occurring less frequently than those in A and B, nevertheless play a significant role in defining the subject area of the materials being read. However, we have yet to develop a method which will precisely indicate subject-laden words in samples of text. Harter's technique represents the current statistical state-of-the-art in this area. The individual semantic significance of subject words remains still a relative matter--relative to the subject being studied, the amount

of material actually sampled, and the degree of precision desired in specifying such words. Luhn (1958), Maron (1961) and Meadow (1973) all deal with this problem in terms of information retrieval and suggest that "peaks" of significance for such words can occur anywhere in the rank frequency list and are often idiosyncratic to particular user needs at specified points in time and the nature of the word frequency distributions being used. Special subject words in the middle and upper part of Sector C would be the best candidates in defining the subject area of the material being read. (However, high frequency general vocabulary words occur in this sector as well). Because the lexical utility of the individual subject words varies considerably, their significance may best by judged by authorities in the subject areas. Words with relatively high frequencies could be termed the most common lexical items, or informing words, while words with low frequencies relate more to highly specialized concepts in the subject.

Sector D. Words in this sector occur relatively infrequently and range from the hapax legomena to frequencies as high as 5 or 6 depending on the size of the corpus. We would assign low significance to such words in developing core vocabularies although such local words in and by themselves may be highly significant in comprehending specific sentences and passages in

text materials.[7] Extremely low frequency words may occur almost completely by chance. These would comprise most of the 12,000 words in the Summers and Edwards corpus which occur only once or twice.

The Summers and Edwards Core Reading Vocabulary contains words selected from Sectors A, B and C of the frequency distribution presented in Table 5. We should point out that there are difficulties in establishing exact cutoff points between sectors in partitioning and considerable subjective decision making is involved in developing core lists. In addition to word frequency data, results from the Kučera and Francis and the Carroll et al. studies are also utilized so as to broaden the frame of reference for the core list and provide some anchor with elementary vocabulary and vocabulary used in adult samples of writing. The list is divided into three parts.

---

[7] It should be noted that although word types represented by just one token in a corpus are given low significance for core vocabulary purposes, the hapax legomena have been found to be of significant value and the magnitude of their presence or absence may even serve to distinguish writings from diverse disciplines (Muller, 1969, pp. 54 and 56). Kučera and Francis (1967, p. vi) report a wide range of one word occurrences across genre and suggest that, ". . . the proportion of hapax legomena is sensitive to some quantitative constraints of corpus size and probably also to qualitative genre differences." This variation in single word occurrence is also apparent in the Summers and Edwards separate subject corpera. The variation in single occurrence word types is also well illustrated in the AH corpus (Carroll, et al, 1971, p. xxxvii).

Part I General Vocabulary. The 87 non-specialty words in Part I all record very high frequencies in Sectors A and B as defined previously. Included are 64 of the highest frequency words from the Summers and Edwards corpus which also record very high frequencies on the Kučera and Francis and the Carroll et al. lists and on at least five of the seven subject lists from the Summers and Edwards corpus. (The first 12 function words on the list account for approximately 27 percent of the total running words in the corpus). An additional 23 function and content words have been included which record relatively high frequencies in the Summers and Edwards corpus, appear on either the Kučera and Francis or the Carroll et al. lists, and also record relatively high frequencies on three of the seven subject lists in the Summers and Edwards corpus. The list is, in essence, an amalgamation of the non-specialty words which occur most frequently across the 469 samples taken from the 37 secondary textbooks (See Table 3). We estimate that the 87 general vocabulary words included in Part I comprise approximately 40 to 50 percent of the total natural language prose in the Summers and Edwards samples. They would also account for a considerable amount of the total token base appearing in the Kučera and Francis and Carroll et al. word lists.

Part II General Subject Vocabulary. The 95 words listed in Part II appear with high frequency in the Summers and Edwards

corpus. These specialty words would fall in the upper part of Sector C. They also appear in both the Kučera and Francis and the Carroll et al. lists, recording frequencies that range from moderate to high. Although some researchers would suggest that it is questionable to attempt to develop high frequency vocabulary lists across a number of subject areas (see Alford, 1971, p. 82, on this point), our core list does include 95 such words. Alford's comments may relate more specifically to highly technical vocabulary (our Part III to some degree) in specialized subject fields while the 95 words included here are extremely basic subject words but also function as more general vocabulary. They represent the very threshold of entry into the lexicon of the subject fields and therefore are essential to a core vocabulary list.

Part III Subject Vocabulary. This section includes seven lists of 414 words which record high frequencies in the subject areas which comprise the Summers and Edwards corpus. These special subject words have high lexical significance and were selected by the authors to provide a representative sampling of words which readers would frequently encounter in the natural language text related to the subject area. All words also appear in the Kučera and Francis and Carroll et al. studies with frequencies varying from moderate to high. The words represent additional very useful entry level subject words.

Part III comprises the most subjective and least generalizable section of the core list. The words selected may be almost completely idiosyncratic to the materials sampled in the Summers and Edwards study. Their best value lies in the fact that they were selected from content that is being used in British Columbia junior secondary grades. The words do correlate well with the lists produced in the Kučera and Francis and the Carroll et al. studies. However, one should not generalize too broadly as to their utility in other locations although the authors feel they have some universality and would most likely be applicable in many situations where a sample of representative subject words is required. The subject textbooks used in British Columbia secondary schools are fairly similar to those used in the United States--particularly in the functional vocabulary contained therein.

SUMMERS AND EDWARDS CORE READING VOCABULARY

Part I General Vocabulary (90% Mastery Criterion)

a	for	it	the
and	in	of	to
are	is	that	you
about	had	not	time
all	has	on	two
also	have	one	up
an	he	only	us
any	him	or	used
as	his	other	very
at	how	see	was
be	I	should	we
been	if	so	were
but	into	some	what

by	its	such	when
called	like	than	which
can	made	their	who
could	many	them	will
do	may	then	with
each	more	there	work
first	most	these	would
from	must	they	your
good	no	this	

Part II General Subject Vocabulary (70% Mastery Criterion)

account	fabric	meal	river
add	family	measure	sales
air	feet	metal	sentence
American	figure	money	set
amount	food	north	small
area	forest	numbers	speed
book	forum	office	steel
British	fuel	oil	surface
business	games	order	table
Canadian	gear	paid	temperature
care	government	paper	tools
cells	heat	part	trade
child	job	pattern	triangle
children	lake	pay	unit
city	land	people	value
copper	language	piston	vegetables
copy	length	point	water
cylinder	light	population	whole
different	line	power	wire
earth	lines	pressure	wood
energy	long	rational	word
engine	made	record	world
English	margin	region	years
example	mass	represent	

PART III Subject Vocabulary (60% Mastery Criterion)

COMMERCE

assignment	credit	gross	salesman
balanced	currency	insurance	savings
bill	customer	journal	selling
business	debt	ledger	shorthand

buys
calculated
carriage
centered
cheque
clerk
compute
consumer
copies
correction
deduct
dictation
divider
due
employee
employer
entries
errors
exemptions
factory
loan
mail
merchandise
operator
parcel
payment
purchases
receipt
receives
salary
spaced
stops
subtract
transactions
type
typewriter
vowels
wage
worker

ENGLISH

adjectives
antecedent
apostrophe
audience
author
auxiliary
awkward
characters
clauses
climax
clues
coherence
colloquial
comedy
communicate
compositions
conclusion
contemporary
critic
define
describing
diary
drama
figurative
formal
grammars
handwriting
homonyms
inflections
interpretation
interrogative
languages
listening
literature
message
metaphor
modifiers
novel
paragraphs
performed
phrase
players
plays
plural
poetry
predicate
presence
prose
reader
rewrite
rhythm
scene
semicolon
singular
slang
speakers
spelled
tale
teach
television
theatre
tone
verbs
vocabulary
writers

HOME ECONOMICS

abdomen
apron
attractive
babies
bakery
bedtime
behavior
beverage
birth
blend
blouses
brand
bread
budget
closet
collar
cook
cottons
crisp
disease
dishes
drawers
dresses
eating
entertaining
fitted
flavored
freezing
labeled
meats
menu
napkins
nursing
nutritious
nylon
oven
packaged
physician
pockets
pork
powders
pregnancy
sew
shelf
show
sink
skirts
slash
sleeves
socks
soiled
strained
suit
tender
textile
texture

bust
buttonhole
calories
canned
cereal
chocolate
grooming
guests
hems
hose
hostess
juices
recipe
rinse
salad
salted
satisfying
saving
utensils
wardrobe
wash
wearing
yarn
yeast

INDUSTRIAL EDUCATION

alloy
artificial
automotive
bearings
blades
boiler
buffer
cable
casting
chisels
connecting
craftsman
curve
drag
equipped
fir
forging
frame
gauge
gears
glued
grained
grooves
hobby
horn
horsepower
joints
layout
locking
machining
mahogany
manufacture
mold
muffler
nails
outlet
pine
pipe
piston
plywood
repair
resin
rocker
sanding
shaping
shops
socket
stroke
thickness
throttle
timing
valves
vapor
velocity
workshop

MATHEMATICS

algebra
binary
congruent
cubes
decimal
denominator
divisible
equals
fraction
geometric
intersect
multiplication
notation
numeral
plus
pyramid
representing
segment
subset

SCIENCE

absorbed
apparatus
astronomer
balloon
beaker
burn
calculate
chemists
chloride
circles
coefficient
compounds
conductor
digestive
earthquake
electromagnet
equalize
experimental
fission
flame
flask
gram
granite
gravity
images
insects
melt
meteorite
microscope
molecular
neutrons
nuclei
organ
photosynthesis
pigments
platinum
protons
radiation
radioactive
scientist
seismic
sensors
solutions
stimuli
substances
telescope
theory
thermometer
tissue
vein
vibrating
wavelength

constant	intestine	react	weight
crystal	investigation	reflecting	
decay	lenses	response	
deposited	lunar	samples	

SOCIAL STUDIES

allies	elected	legislation	resources
arable	empire	majority	revenue
Asia	exploitation	manufactured	route
boom	exploration	medieval	sea
boundaries	feature	metropolitan	seasonal
Caribbean	fertile	migration	settlements
climatic	feudal	neighborhood	site
collective	forage	opposition	society
colonists	freight	overseas	sovereignty
commerce	government	passenger	tariff
commodities	hemisphere	peasant	territory
confederation	highlands	pioneer	transportation
continents	historical	plots	treaty
council	immigrants	prairie	tribes
cultivation	independent	progress	tropical
declaration	institution	rail	urban
depression	invention	rainfall	wealth
domestic	labor	reform	zone
economics	latitude	representation	

The word list incorporates the usual problems associated with developing lists based on graphic form and in automatic computer processing of natural language text. Homonyms are the most obvious of these problems. We cannot assume that all words that are spelled the same represent the same concept. Meaning is context dependent and multiple uses of words like surface and set are not differentiated. However, it was felt that the homonym problem does not do too much injustice to the data, particularly in light of the fact that the major purpose of the

test is to serve as a screening device. It is also customary in word lists to conflate words from the same root (i.e., heat, heating, heated). We have generally not combined such words preferring to let them stand as they appear in the frequency distributions. It can be observed that tokens of one form of a word may be distributed very differently from tokens of another although the stem is the same. Combining them may confound important occurrence frequency differences among words. This procedure follows suggestions made in other key word indexing studies. We have on occasion somewhat juggled word form in correlating the Summers and Edwards corpus with the other two word lists (i.e., as in the case of substituting suit for suits to reflect high frequencies for the singular form in the word lists used for validation). Again, it is felt these changes improve the situation rather than do harm to the data.

The 496 word vocabulary list does provide a representative subset of what we have chosen to refer to as the functional reading vocabulary of secondary print materials. A chief value of the list is its sampling feature. One alternative method in developing such a list would have been to select words, on a logical basis, to represent the total domain of function, content and specialty word types. However, the list would have low curricula validity and it is highly probable that words would be included which do not appear at all, or appear only with very low

frequency, in junior secondary subject materials. The list represents a balanced sample of the lexicon of secondary textbook English.

The list should not be viewed as an instructional program in and by itself although it can be a starting point for instruction. Use of the list can provide very definite answers to the two questions raised in the opening section of this paper. Separate groups of words on the list may have some diagnostic validity, i.e., Does a student have particular difficulty with pronouns and conjunctions as compared with content or specialty words? However, its possible use as a diagnostic tool would have to be substantiated through administration in instructional settings with varied populations. It should also be pointed out that the list is not necessarily a sample of the most common words in spoken English. There are real differences in the way in which the lexicon is used in speech and writing. Reading and spoken vocabulary can be expected to be somewhat parallel but we tend to use a much more restricted and less specialized vocabulary in speaking. (West, 1953; Jones and Wepman, 1966) A productive future project would be development of a screening list reflecting both the spoken and written lexicon in Canadian secondary schools.

Using the Core Reading Vocabulary List in Reading Assessment

Because of the nature of its development, the list has considerable curricula validity and can be used as a criterion referenced assessment, or what the American Psychological Association now refers to as a content referenced test. Because of the recent increase in volume of published literature on the topic, one might assume that criterion-referenced testing is a unique measurement technique related to new developments in education. However, Jaeger (1975) points out that the case for novelty of criterion-referenced testing was dealt a serious blow in 1972 when Airasian and Madaus cited a statement by E. L. Thorndike defining the difference between criterion-referenced and norm-referenced measurements in a 1916 study of the Boston Public Schools. Reports in the more recent literature have compared the advantages and disadvantages of criterion and norm referenced testing and explored theoretical issues and topics such as test construction, item construction, item discrimination, scaling, reliability, and validity in criterion-referenced measurement (Alvord and Brittingham, 1974; Ascher, 1975; Boehm, 1974; Brennan, 1974; Crehan, 1974; Dziuban and Vickery, 1973; Ebel, 1975; Elsner, 1973; Finder, 1974; Fitzgibbon, 1972; Goodwin, 1974; Greco, 1974; Haladyna, 1974; Harsh, 1974; Kifer and Bramble, 1974; Livingston, 1972; Popham, 1974; Roudabush, 1975; Rosen, 1973; and Schwartz, 1974).

Summers

Glaser and Nitko (1971) present perhaps the most widely quoted definition of a criterion referenced test:

> . . . deliberately constructed to yield measurements that are directly interpretable in terms of specified performance standards. Performance standards are generally specified by defining a class or domain of tasks that should be performed by the individual. Measurements are taken on representative samples of tasks drawn from this domain, and such measurements are referenced directly to this domain for each individual measured. (1971, p. 654)

A criterion referenced test may consist of a single task or performance or it may consist of many tasks or performances aimed at measuring many objectives. The ciiterion level may be subjectively or objectively determined but is based on mastery of the specified content. The test tells you what an individual knows or can do while the norm referenced test related performance to a group frame of reference. Criterion referenced tests supplement but do not replace the traditional measuring tasks of norm referenced tests. Ebel (1971) suggests that the essential difference is in the quantitative scale used to express what the individual can do. In norm referencing, the scale is usually anchored in the middle on some average level of performance for a particular group of individuals. In criterion referenced measurement, the scale is anchored at the extremities indicating mastery or non mastery of the behavior. Thus, we can expect performance distributions for criterion referenced tests to be

markedly skewed while those for norm referenced tests would approximate a normal distribution. Housden and Legear (1973) can be consulted for a useful detailed summary comparing major characteristics of the two types of assessment. Criterion referenced tests may be more sensitive to instruction as the behaviors measured have to be precisely defined and the sample content based on specific curricula. Their diagnostic value is thus greater than most norm referenced survey tests. The major difficulty in such testing lies in adequate definition of the criterion behavior to avoid production of what one author has labelled "cloud referenced" tests (1974). Norm referenced tests have their problems but criterion referenced tests do not necessarily correct these. The two types represent differing measuring instruments for different purposes.

In recent years, there has been a surge of interest in criterion referenced measurement in reading, particularly among publishers of commercial tests and materials (Athey, 1975; Davis, 1973, 1974; McDonald and Moorman, 1974; Otto, 1973; Petre, 1973; Pikulski, 1974; Roudabush, 1973; Rude, 1974; Rupley, 1975; Thompson and Dzuiban, 1973; Venezky, 1974; and Womer, 1974). We have extrapolated the following set of criteria for criterion referenced tests from the recent literature on the topic. Good criterion referenced reading tests would include: (1) Clear definition of the reading skill or objective to be

measured.

(2) Precise statement of the acceptable level of mastery.

(3) Mastery levels based on realistically attainable reading goals.

(4) Item content considered to be representative of essential reading behavior.

(5) Tasks to be accomplished that approximate those that take place in real reading situations.

(6) Items arranged in an actual hierarchy or generally accepted sequence of skill development in reading.

(7) Tasks organized so that students can demonstrate mastery of content.

The above criteria have all been considered in establishing the core reading vocabulary test as a criterion referenced assessment of functional vocabulary knowledge.

(1) The reading skill to be measured is recognition of the functional vocabulary needed to read junior secondary instructional materials.

(2) The acceptable level of mastery is set at 90 percent identification of words on Part I; 70 percent identification of words on Part II; and 60 percent identification of words, in any subject area or combination of subject areas, on Part III.

(3) Mastery is based on realistic reading goals. Part I consists of the extremely common vocabulary in English. Lack of mastery indicates poor learning and/or lack of exposure to basic instruction in beginning reading. The Part II vocabulary consists of general subject vocabulary which appears with high frequency in elementary, junior secondary and adult reading materials. Lack of mastery indicates poor learning and/or lack of exposure to reading instruction involving materials at the basic entry level in subject areas. Part III vocabulary represents high frequency vocabulary encountered in specific junior secondary subject areas in B.C. schools, and encountered to some extent in

elementary and adult reading materials as well. Lack of mastery indicates poor learning and/or lack of exposure to basic entry level instructional materials in the secondary subject areas.

(4) The core reading vocabulary is based on a careful sample of representative instructional materials and has been validated through comparison with other word frequency studies. Part I vocabulary is a valid sampling of the crucial common words in English. Parts II and III represent highly frequent words which appear in secondary subjects taught in B.C. junior secondary grades and, to some extent, secondary grades in other locations. The test does include the functional vocabulary needed in reading as judged by the criterion of frequency of occurrence of words in printed materials.

(5) Although actual reading behavior seldom involves responding to just words in isolation, the proficient reader reacts to a large number of words on a sight recognition basis. This is particularly true of functional vocabulary. It is therefore valid to require the reader to react to a recognition task, using words in isolation, to determine the level of sight vocabulary knowledge.

(6) The three part organization of the test reflects the typical sequence of vocabulary development in reading beginning with mastery of the most useful common words and moving to lexical words of varying frequency encountered in the reading materials used in the special subject areas.[8]

---

[8]Statistical validation of a reading skills hierarchy is still open to research although some reasonable sequences, developed through logical analysis and substantiated through conventional wisdom, are quite useful in everyday instruction and curriculum development. It cannot be assumed that the skills measured by reading tests are automatically arranged in a sharp order of simple skills within and leading to more complex skills. However, there is little controversy on the question of vocabulary being a foundation skill. Davis (1972, p. 672) suggests that knowledge of word meanings appears, *a priori*, to be essential for comprehension and is a basic ability in which level of competence exerts great influence on all other skills in comprehension. This assumption would be supported by the work of Bormuth (1969) and others in determining factors related to comprehensibility of print and is also underscored in discussions of readability. For

(7) Students can demonstrate their knowledge by orally identifying words after flash presentation or through identifying known words on a paper and pencil checklist task.

The current literature also suggests that criterion referenced tests serve as diagnostic evaluations and usually lead to follow-up instruction where necessary. In addition, it has also been pointed out that learner expectations should be particularly explicit for such tests. Both conditions are fulfilled by the core reading vocabulary test.

Some further comment should be added with respect to (2) above. As Otto (1973) points out, determining proficiency standards for criterion referenced tests can be troublesome. Word frequency data was a prime factor in developing the rationale in establishing the percentage cutting scores in calibrating the test. It was felt that the student should exhibit almost total sight mastery, give or take a few words, of the most commonly used words in the language. Thus, the 90 percent criterion established for Part I. Correspondingly lower percentages of mastery are required for less frequently occurring words, words with low repeat rates, and for those words which are more specific to subject areas where the probability of previous exposure to such vocabulary in basic reading instruction is

---

example. Klare (1969) suggests that, in fact, word frequency encompasses most other variables related to readability. The word factor has probably received more study than any other variable in readability research. (Bailey, 1973, 1975)

lower. Thus, the more liberal 70 percent criterion for Part II and the 60 percent criterion for Part III. Part III has been organized so that users may assess knowledge in any one or combination of subject areas depending on the specific purposes in testing. In addition, criteria for Parts II and III are deliberately liberal to allow for the fact that the reader is responding to an artificial situation and has little of the recognition information--context and graphic aids--that he would normally rely on to extract meaning from connected discourse in prose passages.

The core reading vocabulary is best used as an aid in determining general reading achievement and establishing a rough starting point for instruction. Various common testing techniques can be adapted in using the list. Perhaps the most functional method of testing is to reproduce each word on a separate card. Introduce the task to the student and ask him to read each word aloud. Allow only a brief exposure (up to 5 seconds). A simple checklist can be used to record errors. Potentially useful diagnostic information can be obtained by carefully noting errors, mispronunciations, use of oral language, and the general method of attack used on difficult words. Later returning to words and encouraging the student to try and work out difficult words can provide additional insight into his word attack techniques. Although the objective is to determine recognition

vocabulary, a rough estimate of comprehension can be made by using selected words and asking the student to "use the word in a sentence" or to briefly "tell me what the word means". Part I and II should be administered first. If the student is having excessive difficulty, there is no need to continue to the specialized vocabulary in Part III. For Part III, the lists can be used separately for each subject area, sampled across subject areas, or used in their entirety. The card approach can also be varied using a simple listing of the words.

Another testing technique is to provide the student with the words set up as a self-inventory checklist. The student is asked to read the list and check the words he doesn't know. To check validity, randomly select a number of words checked as known and ask the student to pronounce the words and use them in a sentence. Students can also be asked to differentiate their responses and indicate how well they know the word by using various marks to indicate: "I've never heard or seen the word," "I know the word somewhat," and, "I know it well and use it." This technique provides good data on reading recognition but lacks the advantage of also checking oral usage of the words. However, the technique is very useful for group testing. (Dale & O'Rouke, 1971)

A third technique is to organize the words for each part

of the list into random groups of four or five, type them horizontally on separate lines, and double space between lines as follows:

1. I they is two us
2. account business measure land river
3. vocabulary meats theory tissue economics

Call out one word in the line and have the student check the word called out. Once the lists have been organized in this fashion, innumerable forms are possible by varying the word selected from each line. Again, this is a useful technique for group testing. A variation is to provide a broader meaning context by calling the word, using it in a sentence, and calling it again. We have found it useful to develop a tape for testing. A tape presents the verbal stimuli uniformly and allows for performance comparisons across groups. High test reliability is also possible if good equipment is used.

Another useful testing technique consists of building short passages of about 100 words which are heavily loaded with words from the sections of the test. The student is allowed to first read the passage silently and then asked to read it aloud while the examiner scores errors. Some of the more classic testing modes (i.e., multiple choice: completion, definition and paraphrase) are useful in estimating content achievement after instruction but are less appropriate for this type of screening

testing, particularly when the target is level of recognition vocabulary.

The three parts of the test also lend themselves to some categorization of stages of language proficiency. The test helps answer the question, What is the student capable of doing as a result of cumulative learning experience? Language is learned moving from listening to speaking, reading and writing. Vocabulary merits separate testing because it is the important element shared by all four skill areas. Factor analytic studies consistently demonstrate that comprehension is cognitively oriented and consists of knowledge of word meanings and ability to reason in verbal terms. A satisfactory level of test performance would indicate the student is functioning at approximately the beginning intermediate stage (in terms of functional vocabulary) for second language learners and at around the ending intermediate level/ beginning junior secondary level for native speakers.[9] In general, criterion referenced tests require longer time limits

---

[9] A rough model of stages of reading development would include the acquisition stage (approximately the primary grades) and the application stage (the intermediate grades and beyond). Movement from the acquisition to the application stage represents the broadening of reading ability into the expository patterns represented by the writing in the subject areas. Satisfactory performance on the screening test roughly indicates that both second language students and native speakers possess a fairly adequate base of functional recognition vocabulary and are well into the application stage in reading development.

for administration than the norm referenced type. Parts I and II could comprise one sitting and Part III a second. Administration should be open ended but an anticipated time limit would range from 45 to 60 minutes per sitting depending on the mode of testing and the number of words covered. Proportionately longer time periods would be needed for oral testing.

Some Concluding Remarks

This paper has described a model for the development of a screening test based on the functional recognition vocabulary which is generic to secondary textbook materials. The model can be used in other research on word list development. Words occur in patterns in text samples and word occurrence frequencies can be conceived of as variables that can be manipulated and that are likely to behave in probabilistic fashion. Occurrence frequencies can also be used in partitioning word lists as an aid in forming core vocabularies. Words with very high occurrence frequencies, distributed non-uniformly across subject samples irrespective of subject content, constitute the very essential common vocabulary in reading. Those words which tend to cluster in samples from one subject area, but also distribute to some extent across other subject samples, comprise the general subject vocabulary. Words which tend to cluster only in samples related to a particular area form the special subject vocabulary. Words with very low occurrence frequencies which distribute minimally

across samples, if at all, constitute rare vocabulary and are not candidates for core vocabulary lists. Since redundancy is clearly related to comprehensibility, repeat rate frequency of words also becomes an important variable in core vocabulary selection and instructional materials development.

An approximate calibrating system for the word list has been proposed based on the current literature related to criterion-referenced testing. However, no validating research yet supports the proposed scoring scheme and it is presented as a speculative venture in the hope of stimulating further definitive research on use of such schema. It should also be borne in mind that adequate recognition vocabulary is a necessary but not sufficient condition for word knowledge and comprehension. The screening test provides data on the words a student may recognize, and perhaps also understand, but exact word knowledge and comprehension would have to be substantiated through additional checks. The test should aid in screening those who are prepared for secondary textbook reading from those who are not.

Finally, it goes almost without saying that it is important for educators to increase their linguistic sensitivity and carefully consider the language demands of textbooks, particularly those which serve as the initial entry into the important concepts and ideas of secondary subject areas. In the near future, it may be possible for publishers to provide, as a matter of course, word

frequency distributions, repeat rate statistics, and other data on all materials recommended for use in school programs as a by-product in printing. Such data would help considerably in testing, curriculum development, research and teaching. Current technology is quite up to the task. Just as cataloguing in press is becoming more common in library materials, arrangements could be made for type and token characteristics to accompany educational materials. Photo composition methods, computer based typesetting, man-machine interaction, systems time sharing and the like are all applicable to book production and such data could be an easily obtainable by-product (Law, 1974). The role of the author and publisher could sensibly mesh more closely with the consumer (the teacher and student) in the conception, production and use of school instructional materials. Terrant (1975) suggests that we may now be on the verge of the long awaited shift to broader use of computers in publishing.

> For years we have heard about the magic blending of computer technology and information processing as a solution to our publishing ills, and usually we are disappointed. But in the past few years, many of the long-promised capabilities have materialized. We have seen increased use of the computer, word-processing systems, and photo composition. The range of versatility of both hardware and software has increased dramatically, usually as unit hardware operating costs have decreased. (p. 295)

References

Airasian, P. W. and Madaus, G. F. Criterion-referenced testing in the classroom. Measurement in Education, May, 1972.

Alford, M. H. T. Computer assistance in language learning and in authorship identification. In R. A. Wisbey (Ed.) The computer in literary and linguistic research. London: Cambridge University Press, 1971, pp. 77-92.

Alvord, D. J. and Brittingham, B. E. Evaluating performance on national assessment objectives: norm-referenced and criterion-referenced interpretations. Journal of Educational Research, 1974, 68, 59-61.

Ascher, G. Some nonparametric approaches to the use of criterion-referenced statewide test results in the evaluation of local district education programs. Paper presented at the Annual Meeting of the American Educational Research Association, 1975. (ED 106 369)

ASIS Cumulative index to the annual review of information science and technology Volumes 1-10. Washington: American Society for Information Science, 1975.

Athey, I. and O'Reilly, R. A criterion-referenced testing model for assessing growth in reading. Paper presented at the annual meeting of the National Council for Measurement in Education, 1975. (ED 105 422)

Bailey, S. D. Recent trends and developments in research involving the cloze procedure. Unpublished manuscript. University of British Columbia, 1973.

Bailey, S. D. The application of selected principles of sentence rhetoric to the study of the readability of instructional materials: a conceptual base for research and investigation. Unpublished masters paper. University of British Columbia, 1975.

Boehm, A. E. Criterion-referenced assessment for the teacher. Teachers College Record, 1973, 75, 17-26.

Bonwit, K. and Aste-Tonsmann, J. Negative dictionaries. Information Storage and Retrieval (ISR XVIII), Ithaca, New York: Department of Computer Science, Cornell University, 1970.

Bookstein, A. and Swanson, D. R. Probabilistic models for automatic indexing. Journal of the American Society for Information Science, 1974, 25, 312-318.

Bookstein, A. and Swanson, D. R. A decision theoretic foundation for indexing. Journal of the American Society for Information Science, 1975, 26, 45-50.

Bormuth, J. R. Development of readability analyses (Final report U.S. Office of Education Project No. 7-0052), University of Chicago, 1969.

Brennan, R. L. Some potential uses of decision-thoretic (confidence) testing in the analysis of criterion-referenced item data. Washington: Educational Resources Information Center, 1974. (ED 093 993)

British Columbia Department of Education. Prescribed textbooks 1972-73 Grades I-XII. Victoria: Curriculum Development Branch, 1972.

Buckingham, B. R. and Dolch, W. E. A combined word list. Boston: Ginn & Co., 1936.

Carroll, J. B. Fundamental considerations in testing for English language proficiency of foreign students. Testing the English proficiency of foreign students. Washington: Center for Applied Linguistics, 1961.

Carroll, J. B. Behind the scenes in the making of a corpus-based dictionary and a word frequency book. Paper presented at the meeting of the National Council of Teachers of English, 1971. (ED 056 842)

Carroll, J. B., Davies, P., and Richman, B. Word frequency book. Boston: American Heritage Publishing Co., Inc., and Houghton Mifflin Company, 1971.

Crehan, K. D. Item analysis for teacher-made mastery tests. Journal of Educational Measurement, 1974, 11, 255-62.

Curtice, R. M. and Jones, P. E. Distribution constraints and the automatic selection of an indexing vocabulary. Proceedings of the American Documentation Institute Annual Meeting, 1967, 4, 152-156.

Dale, E. and Chall, J. S. A formula for predicting readability. Educational Research Bulletin, 1948, 27, 11-20, 37-54.

Dale, E. and O'Rourke, J. Techniques of teaching vocabulary. Palo Alto, California: Field Enterprises, Inc., 1971.

Dauzat, S. V. Structure word usage in the verbal discourse of two groups of children. Unpublished doctoral dissertation. University of Mississippi, Oxford, 1968.

Davies, P. New views of lexicon. In J. B. Carroll, P. Davies and B. Richman Word frequency book. Boston: American Heritage Publishing Co., Inc., and Houghton Mifflin Company, 1971.

Davis, F. B. Psychometric research on comprehension in reading. Reading Research Quarterly, 1972, VII, 628-678.

Davis, F. B. Criterion-referenced measurement. Princeton: Educational Testing Service, 1973. (ED 073 143)

Davis, F. B. Criterion-referenced tests: a critique. In W. E. Blanton (Ed.) Measuring reading performance. Newark, Delaware: International Reading Association, 1974.

Dennis, S. F. The construction of a thesaurus automatically from a sample of text. In Mary E. Stevens et al. (Eds.) Statistic Association Methods for Mechanized Documentation. Washington: National Bureau of Standards Miscellaneous Publication 269, 1965.

Dolch, E. W. A basic sight vocabulary. Elementary School Journal, 1936, 36, 456-460.

Doležel L. and Bailey, R. W. (Eds.) Statistics and Style. New York: American Elsevier Publishing Co., Inc., 1969.

Dunn, M. K. and Laffey, J. L. Research on reading: word lists. Washington: Educational Resources Information Center, 1969. (ED 030 778)

Durr, W. K. Computer study of high frequency words in popular trade juveniles. The Reading Teacher, 1973, 27, 37-42.

Dziuban, C. D. and Vickery, K. W. Criterion-referenced measurement: Some recent developments. Educational Leadership, 1973, 30, 483-486.

Ebel, R. L. Evaluation and educational objectives. Journal of Educational Measurement, 1973, 10, 273-279.

Ebel, R. L. Educational tests: valid? biased? useful? Phi Delta Kappan, 1975, 57, 83-89.

Edwards, P. A computer generated corpus and lexical analysis of English language instructional materials prescribed for use in British Columbia junior secondary grades. Unpublished doctoral dissertation. University of British Columbia, Vancouver, 1974.

Elsner, P. J. Criterion referenced assessment and its classroom use as viewed by teachers. School Science and Mathematics, 1973, 9, 730-734.

Eskey, D. E. A new technique for the teaching of reading to advanced students. Paper given at the fourth annual TESOL Convention, 1970. (ED 038 649)

Finder, M. Criterion tests and standardized tests. Paper presented at the annual meeting of the New York State English Council, 1974. (ED 092 957)

Finocchiaro, M. English as a second language: from theory to practice. New York: Regents Publishing Co., 1964.

Fitzgibbon, T. J. Norm referenced and criterion referenced tests from a publisher's point of view. Paper presented at the Convention of the American Psychological Association, 1972. (ED 069 787)

Fries, C. C. American English grammar. New York: Appleton Century Co., 1940, pp. 110-114.

Fries, C. C. The structure of English. New York: Harcourt, Brace & Co., 1952.

Fry, E. Basic vocabulary--the instant words. Reading instruction for classroom and clinic. New York: McGraw Hill, 1972.

Glaser, R. and Nitko, A. Measurement in learning and instruction. In R. L. Thorndike (Ed.) Educational measurement. Washington: American Council on Education, 1971.

Goodwin, C. R. The development of a diagnostic test for determining secondary students' knowledge and understanding of certain basic language concepts. Unpublished doctoral dissertation, Arizona State University, 1974.

Greco, T. H. Is there really a difference between CR and NR measurements? Educational Technology, 1974, 14, 22-25.

Haladyna, T. M. An investigation of full and subscale reliabilities of criterion-referenced tests. Paper presented at the annual meeting of the American Educational Research Association, 1974. (ED 091 435)

Hambleton, R. K. and Novick, M. R. Towards an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, 159-70.

Harris, A. J. and Jacobson, M. D. Basic elementary reading vocabularies. London: McMillan and Co., 1972.

Harris, David P. Testing English as a second language. New York: McGraw-Hill Book Co., 1969.

Harsh, J. R. Trees, branches and leaves, revisited-norm, domain, objective and criterion-referenced assessments for educational assessment and evaluation. AMEG Monograph No. 1. Fullterton, California: California Personnel and Guidance Association, 1974. (ED 097 344)

Harter, S. P. A probabilistic approach to automatic keyword indexing: Part I, on the distribution of specialty words in a technical literature. Journal of the American Society for Information Science, 1975, 26, 197-206.

Harter, S. P. A probabilistic approach to automatic keyword indexing: Part II, an algorithm for probabilistic indexing. Journal of the American Society for Information Science, 1975, 26, 283-289.

Herdan, G. Language as Choice and Chance. Groningen: P. Noordhoff, 1956, p. 32.

Hill, L. A. Selected articles on the teaching of English as a foreign language. London: Oxford University Press, 1969.

Horn, E. A basic writing vocabulary. Iowa City: University of Iowa Monographs in Education, First Series, No. 4, April 1926.

Housden, J. L. and Legear, L. An emerging model: criterion referenced evaluation. Thrust for Educational Leadership, 1973, 2, 41-6.

Jaeger, R. M. Some new developments and discoveries for evaluative analysis. Center on Evaluation Development and Research Quarterly, 1975, 9, 3-7.

Johnson, D. B. Computer frequency control of vocabulary in language learning reading materials. Instructional Science, 1972, 1, 121-131.

Jones, Lyle V. and Wepman, J. M. A spoken word count. Chicago: Language Research Associaties, 1966.

Kelly, L. G. English as a second language: an historical sketch. English Language Teaching, 1971, 25, 120-132.

Kifer, E. and Bramble, W. The calibration of a criterion-referenced test. Paper presented at the Annual Meeting of the American Educational Research Association, 1974. (ED 091 434)

Kitchin, A. T. Readings: English as a second language, Books I, II. New York: Reader's Digest Services, Inc., 1964.

Klare, G. R. The role of word frequency in readability research. Elementary English, 1968, 45, 12-22.

Kochen, M. Principles of information retrieval. Los Angeles: Melville Publishing Co., 1974.

Kučera, H. Computers in language analysis and in lexicography. American heritage dictionary of the English language. Boston: American Heritage Publishing Co., Inc., and Houghton Mifflin Company, 1969.

Kučera, H. and Francis, N. W. Computational Analysis of Present-Day American English. Providence, Rhode Island: Brown University Press, 1967.

Lancaster, F. W. Information retrieval systems. New York: John Wiley & Sons, 1968.

Lancaster, F. W. Vocabulary control for information retrieval. Washington: Information Resources Press, 1972.

Law, T. G., Jr. Cold type composition: its impact on library and information science. Journal of the American Society for Information Science, 1974, 25, 319-326.

Livingston, S. A. Criterion referenced applications of classical test theory. Journal of Educational Measurement, 1972, 9, 13-26.

Luhn, H. P. The automatic creation of literature abstracts. IBM Journal of Research and Development, 1958, 2, 159-165. Reprinted in A. W. Elias (Ed.) Key papers in information science. Washington: American Society for Information Science, 1972, pp. 87-93.

McDonald, T. F. and Moorman, G. B. Criterion referenced testing for functional literacy. Journal of Reading, 1974, 17, 363-366.

Mandelbrot, B. An informational theory of the statistical structure of language. In Willis Jackson (Ed.) Communication Theory. London: Butterworths Scientific Publications, 1953, pp. 486-502.

Mandelbrot, B. Information theory and psycholinguistics: a theory of word frequencies. In P. F. Lazarfield and N. W. Henry (Eds.) Readings in Mathematical Social Science. Chicago: Science Research Associates, 1966.

Maron, M. E. Automatic indexing: an experimental inquiry. Journal of the Association for Computing Machinery, 1961, 8, 404-417. Reprinted in A. W. Elias (Ed.) Key papers in information science. Washington: American Society for Information Science, 1972, pp. 94-107.

Meadow, C. T. Statistical selection of index terms. Los Angeles: Melville Publishing Co., 1973.

Meyers, W. E. A study of usage items. College Composition and Communication, 1972, 23, 55-69.

Miller, A. Computer analysis of Dr. P. Edwards' grade eight, nine and ten (B.C.) language study corpus. Unpublished manuscript. U.B.C. Computing Centre, 1974.

Montgomery, C. A. Automatic language processing. In C. A. Cuadra (Ed.) Annual Review of Information Science and Technology Vol IV. Chicago: Encyclopaedia Britannica, 1969, pp. 146-74.

Muller, C. Lexical distribution reconsidered. In L. Doležel and Richard W. Bailey (Eds.) Statistics and Style. New York: American Elsevier Publishing Co., Inc., 1969.

Nisbet, J. D. Frequency counts and their uses. Educational Research, 1960, 3, 51-64.

Otto, W. Evaluating instruments for assessing needs and growth in reading. In Walter MacGinitie (Ed.) Assessment Problems in Reading. Newark, Delaware: International Reading Association, 1973, pp. 14-20.

Pennock, C. D. Quick word lists for Canadian readers. Alberta Journal of Educational Research, 1974, XX, 8-14.

Petre, R. M. The development of criterion-referenced tests in reading. Charleston, West Virginia: West Virginia State Department of Education, 1973. (ED 078 048)

Pikulski, J. J. Criterion referenced measures for clinical evaluations. Paper presented at the Annual Meeting of the College Reading Association, 1973. (ED 085 660)

Popham, J. W. An approaching peril: cloud referenced tests. Phi Delta Kappan, 1974, 55, 614-615.

Roberts, A. H. A Statistical Linguistic Analysis of American English. London: Mouton & Co., The Hague, 1965.

Roudabush, G. E. Item selection for criterion-referenced tests. Paper presented at annual meeting of the American Educational Research Association, 1973. (ED 074 147)

Rude, R. T. Objective-based reading systems: an evaluation. Reading Teacher, 1974, 28, 69-75.

Rupley, W. H. ERIC/RCS review: criterion-referenced testing. The Reading Teacher, 1975, 28, 426-27, 429.

Salton G. Information analysis and dictionary construction. Automatic Information Organization and Retrieval. New York: McGraw Hill Book Co., 1968.

Salton, G., Yang, C. S., and Yu, C. T. A theory of term importance in automatic text analysis. Journal of the American Society for Information Science, 1975, 26, 33-44.

Schwartz, H. P. Testing for instructional purposes: norm referenced-criterion referenced. Paper presented at the State Convention of the Kansas Personnel and Guidance Association, 1974. (ED 096 335)

Siliakus, H. J. Computer-aided word research. Babel, 1967, 3, 19-21.

Sparck Jones, K. Automatic keyword classifications. London: Butterworths, 1971.

Stone, D. C. and Rubinoff, M. Statistical generation of a technical vocabulary. American Documentation, 1968, 19, 411-412.

Svenonius, E. An experiment in index term frequency. Journal of the American Society for Information Science, 1972, 23, 109-121.

Terrant, S. W. The computer and publishing. In C. A. Cuadra and A. W. Luke (Eds.) Annual Review of Information Science and Technology Vol. 10. Washington: American Society for Information Science, 1975, pp. 274-301.

Thompson, R. and Dziuban, C. Criterion referenced tests in perspective. The Reading Teacher, 1973, 27, 292-94.

Thonis, E. W. Teaching reading to non-English speakers. New York: The Macmillan Co., 1970.

Thorndike, E. L. A teachers' word book of 10,000 words. New York: Teachers College, Columbia University, 1921.

Thorndike, E. L. and Lorge, I. The teachers' word book of 30,000 words. New York: Teachers College, Columbia University, 1944.

Toth, W. Review of recent literature and conceptual analysis on criterion referenced testing in reading. Unpublished manuscript. University of British Columbia, 1975.

Underwood, B. J. Word recognition memory and frequency information. Journal of Experimental Psychology, 1972, 94, 276-83.

Venezky, R. L. Testing in reading: assessment and instructional decision making. Urbana, Illinois: National Council of Teachers of English, 1974.

West, M. A general service list of English words. London: Longmans', Green & Co., Ltd., 1953.

Womer, F. B. What is criterion referenced measurement? In W. F. Blanton (Ed.) Measuring Reading Performance. Newark, Delaware: International Reading Association, 1974.

Wyllys, R. W. Measuring scientific prose with rank-frequency ("Zipf") curves: a new use for an old phenomenon. In Charles W. Husbands and R. L. Tighe (Eds.) Information Revolution: Proceedings of the 38th ASIS Annual Meeting Vol. 12. Washington: American Society for Information Science, 1975, pp. 30-31. (Also, personal conversation with the author, ASIS Conference, Boston, 1975).

Young, C. E. Development of language analysis procedures with application to automatic indexing. Columbus, Ohio: Computer and Information Science Research Center, 1973.

Yule, G. V. The statistical study of literary vocabulary. London: Cambridge University Press, 1944.

Zipf, G. K. The meaning-frequency relationship of words. The Journal of General Psychology, 1945, 33, 251-256.

Zipf, G. K. Human Behavior and the Principal of Least Effort. Cambridge, Mass.: Addison-Wesley Publishing Co., 1949, 19-55.

Zipf, G. K. The Psycho-biology of Language: An Introduction to Dynamic Philology. Boston: Houghton Mifflin; reprint edition, Cambridge, Mass.: MIT Press, 1965.

Textbooks Sampled

Commerce (41 samples)

1. Wanous, S. J. Personal typewriting. Toronto: W. J. Gage Ltd., 1967. (11 samples)

2. Trotter, C. A., and Glover, P. C. The junior clerk. Toronto: Sir Isaac Pitman (Canada) Ltd., 1970. (14 samples)

3. No author. New basic course in Pitman Shorthand. Toronto: Sir Isaac Pitman (Canada) Ltd., 1964. (7 samples)

4. Dame, J. Frank et al. Exploring business. Toronto: McGraw-Hill Co., of Canada Ltd., 1968. (9 samples)

English (80 samples)

5. McMaster, R. J. The craft of writing. Don Mills, Ontario: Longman Canada Ltd., 1965. (7 samples)

6. Newell, L. H. and MacDonald (eds.) Short stories of distinction. Agincourt: The Book Society of Canada Ltd., 1960. (10 samples)

7. Penner, G. Philip, and McConnell, Ruth E. Learning English. Toronto: The MacMillan Company of Canada Ltd., 1963. (20 samples)

8. Dawe, Alan, and Gibbons, Maurice. The accomplished reader. Don Mills, Ontario: Bellhaven House, 1964. (7 samples)

9. Bruyn, Jan de. Prose Readings. Ontario: Longmans Canada Ltd., 1964. (10 samples)

10. Bullocke, J. G. (ed.) The Harrap book of modern short stories. Toronto: Clarke, Irwin and Co. Ltd., 1964. (10 samples)

11. Ross, Malcolm and Stevens, John (eds.) Eighteen stories. Don Mills, Ontario: J. M. Dent and Sons (Canada), 1965. (12 samples)

12. Voaden, Herman. Drama IV. Toronto: The MacMillan Co. of Canada Ltd., 1965. (4 samples)

Home Economics (98 samples)

13. Barclay, M. S. and Champion, F. Teen guide to homemaking. Toronto: McGraw-Hill Co. of Canada Ltd., 1968. (22 samples)

14. Shank, D. E., et al. Guide to modern meals. Toronto: McGraw-Hill Co. of Canada Ltd., 1970. (21 samples)

15. Todd, E., and Roberts, F. Clothes for teens. Toronto: D. C. Heath Canada Ltd., 1970. (22 samples)

16. Shuey, R. M., et al. Learning about children. Philadelphia: J. B. Lippincott Co., 1964. (14 samples)

17. No author. Up the years from 1 to 6. Department of National Health and Welfare, Ottawa, Canada, 1967. (10 samples)

18. Duffie, M. A. So you are ready to cook. Minneapolis: Burgess Publishing Co., 1964. (9 samples)

Industrial Education (63 samples)

19. Laustrup, Jes et al. Exploring industrial education. New York: McGraw-Hill Co. of Canada Ltd., 1968. (9 samples)

20. Groneman, Chris H. General woodworking. Toronto: McGraw-Hill Co. of Canada Ltd., 1965. (13 samples)

21. Feirer, John L. General metals. Toronto: McGraw-Hill Co. of Canada Ltd., 1965. (16 samples)

22. Worthington, Robert M. General power mechanics. Toronto: McGraw-Hill Co. of Canada Ltd., 1970. (25 samples)

Mathematics (35 samples)

23. Brumfiel, et al. Introduction to mathematics. Reading, Mass.: Addison-Wesley Publishing Co. Inc., 1962. (14 samples)

24. Eicholy, R. E., et al. Modern general mathematics. Don Mills, Ontario: Addison-Wesley (Canada) Ltd., 1966. (7 samples)

25. Wilcox, M. S., and Yarnelle, J. E. Mathematics: a modern approach. Don Mills, Ontario: Addison-Wesley (Canada) Ltd., 1966. (14 samples)

Science (75 samples)

26. Cannon, G. H., et al. Labtext in science: Book 1. Toronto: The Copp Clark Publishing Co., 1968. (9 samples)

27. Anastasiou, Clifford J. et al. Reading about science 1. Holt, Rinehart and Winston of Canada, Ltd., 1968. (11 samples)

28. Rasmussen, W. H., and Schmid, M. C. Developing science concepts in the laboratory. Scarborough, Ontario: Prentice-Hall of Canada Ltd., 1968. (13 samples)

29. Forster, M., et al. Reading about science 2. Holt, Rinehart and Winston of Canada Ltd., 1969. (11 samples)

30. Schmid, M. C. (ed.) Extending science concepts in the laboratory. Scarborough, Ontario: Prentice-Hall of Canada Ltd., 1970. (17 samples)

31. Woodrow, J. Reading about science 3. Toronto: Holt, Rinehart and Winston of Canada, Ltd., 1970. (14 samples)

Social Studies (77 samples)

32. Carswell, Bordon E. Man in the tropics. Scarborough, Ontario: Bellhaven House Ltd., 1968. (15 samples)

33. Williams, Geoffrey. The shaping of modern Europe. Toronto: The MacMillan Company of Canada Ltd., 1968. (7 samples)

34. Carswell, G. E., et al. Man in the great community. Scarborough, Ontario: Bellhaven House Ltd., 1969. (8 samples)

35. Innis, Hugh R. Our world of change. Toronto: McGraw-Hill Company of Canada, Ltd., 1969. (5 samples)

36. Tomkins, G. S. A regional geography of North America. Toronto: Gage Educational Publishing Ltd., 1970. (30 samples)

37. Lower, J. A. A nation developing: a brief history of Canada. 1970. (12 samples)